Can Mathematics Make Music Sound Good?

Tuning Systems as Optimisation Problems

William Odumosu

First Year BSc Mathematics, Department of Mathematics, University of Manchester

The problem: what does "optimal" mean?

Twelve-tone equal temperament (12-TET) is commonly described as the "best compromise" for musical tuning. It enables unrestricted modulation and transpositional invariance at the cost of harmonic impurity.

But this claim is typically qualitative. It is unclear what is being optimised, which intervals are prioritised, or which loss function is implicitly assumed.

The claim

12-TET is the "best compromise"

The gap

Best under which objective? Nobody specifies.

This talk

Formulate tuning choice as an explicit optimisation problem.

Consonance, frequency ratios, and the cent

Consonant intervals correlate with low-integer frequency ratios, motivated by alignment of partials in the harmonic series. Fixed-pitch instruments require a discrete pitch inventory; some intervals must deviate from just targets to satisfy **octave closure, modulatory freedom, and transpositional invariance.**

Interval	Ratio	Just cents
Perfect fifth	3 : 2	702.0¢
Major third	5 : 4	386.3¢
Perfect fourth	4 : 3	498.0¢
Minor third	6 : 5	315.6¢

Logarithmic pitch: $c = 1200 \log_2(f_2 / f_1)$

One octave = 1200 cents. One n-TET step = 1200/n cents. Octave shifts become additive translations in log space.

12-TET divides the octave into 12 equal steps of 100¢ each. This maximal symmetry means **every key sounds identical** (transpositional invariance), but **no interval is perfectly in tune**. Every consonance deviates from its just target. The question is: by how much, and can we do better?

The optimisation framework

For step size $s_\square = 1200/n$: $k^{(n)}=\mathrm{round}(c^{just}/s_\square)$, $c^{(n)}=k^{(n)}\cdot s_\square$, $e^{(n)}=|c^{(n)}- c^{just}|$

Three aggregation norms:

L^1

$E_1 = \Sigma\ w_i|e_i|$

Total impurity.
Less sensitive
to outliers.

L^2

$E_2 = \Sigma\ w_i(e_i)^2$

Penalises large
deviations. Hates
outliers.

$L\infty$

$E\infty = \max|e_i|$

Worst interval
determines
the score.

Four modelling choices define the objective: target set (which intervals?), error measure (absolute cents deviation), aggregation norm (L^1, L^2, $L\infty$), and weighting scheme (which intervals matter most?).

"12-TET is optimal" has no mathematical content unless all four are specified. Change any one, and the ranking of temperaments can change.

Weighting schemes: what does the music care about?

Weights $w_i > 0$ encode harmonic priorities. The choice is not arbitrary: it reflects what the music demands.

Uniform

`wᵢ = 1 for all`

Baseline: all intervals equally important. No musical context assumed. Useful as a neutral benchmark, but rarely reflects how music actually works.

Fifth-prioritised

`w(P5)=w(P4)=4`

Reflects tuning traditions built around the circle of fifths. Relevant for medieval polyphony and any tradition where open fifths dominate the texture.

Third-prioritised

`w(M3)=w(m3)=4`

Prioritises the intervals most responsible for chord quality. Relevant for Renaissance and Romantic music, where rich triadic harmony is central.

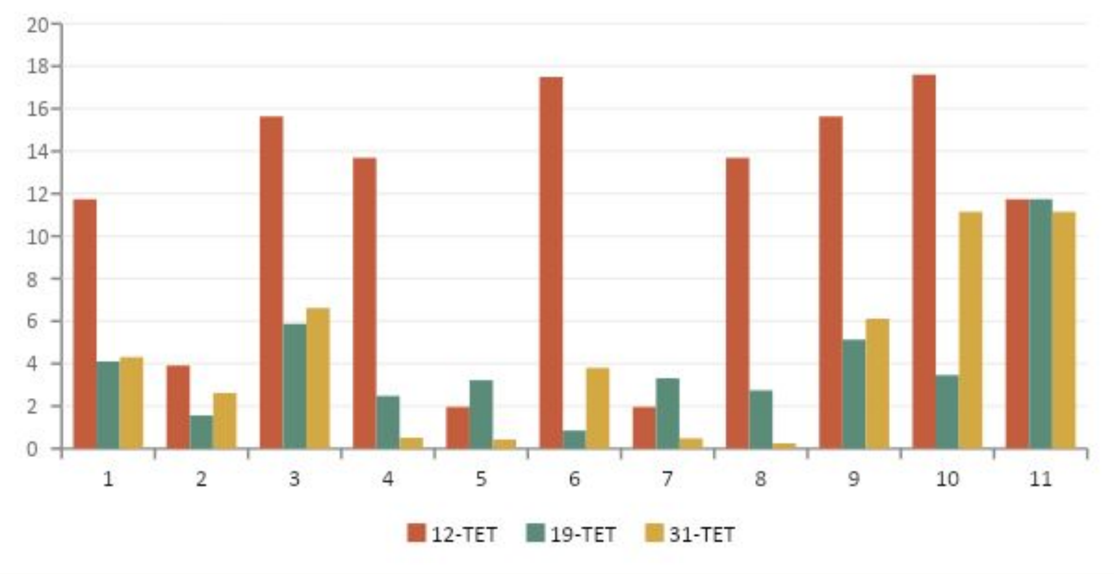
Triadic

`w=3 for triadic set`

Weights all six intervals that build major and minor triads. The most comprehensive scheme for music centred on chordal harmony (Baroque through modern pop).

The weighting scheme transforms "what sounds good" from a vague aesthetic preference into a measurable mathematical parameter.

Computed results



Aggregate errors

TET	Scheme	E_1	E_2	E_∞
12	uniform	125.0	1777.6	17.6
12	thirds	213.0	3073.5	17.6
19	uniform	94.2	1095.1	14.6
19	thirds	116.8	1258.0	14.6
31	uniform	56.0	399.8	11.1
31	thirds	76.3	508.3	11.1

> 12-TET's major third is 13.7¢ off; 31-TET achieves 0.5¢. But 31-TET's minor seventh is worse. No temperament dominates across all intervals.
>
> 31-TET substantially outperforms 12-TET under every weighting scheme tested. But the magnitude of the gap depends on the objective. The correct claim is always conditional: "12-TET is optimal under objective O."

Beyond equal temperament

Quarter-comma meantone

Tempers each fifth by ¼ of the syntonic comma:

$$\Sigma = 1200 \log_2(81/80) \approx 21.51¢$$

$$c^{MT} = 701.96 - \Sigma/4 \approx 696.58¢$$

Four meantone fifths minus two octaves yield an **exactly pure major third:**

$$4c^{MT} - 2400 = 1200 \log_2(5/4) = 386.31¢$$

Cost: 12 meantone fifths don't close the octave. The ≈41.1¢ discrepancy produces a **wolf interval** in one key.

Error shifts from "everywhere a little" to "most places pure, one place catastrophic."

Well temperaments

Circulating temperaments that avoid a single catastrophic wolf while retaining unequal key colour.

All keys are usable, but not equivalent. Different keys have distinct harmonic character from systematic differences in third and fifth sizes.

Historical context: Bach's Well-Tempered Clavier almost certainly implies a circulating unequal temperament, not 12-TET. The exact tuning is debated, but the musical point is robust: unequal temperaments produce audible key colour.

In optimisation terms: controlled unevenness. Distribute error unequally to preserve key character.

Repertoire-linked interpretation

Weights can be induced from a piece rather than chosen arbitrarily. The composer implicitly defines the objective function.

1. Select a piece with a clear harmonic plan
2. Compute chord tokens from the score
3. Map each chord to its interval multiset
4. Count interval frequencies → weights
5. Evaluate temperaments under E_1, E_2, E_∞

Example: `Major triad → {M3, P5}` `Minor triad → {m3, P5}` `Dom. 7th → {M3, P5, m7}`

A Bach prelude heavy in major triads will weight thirds and fifths strongly. A Debussy piece built on whole-tone movement will weight entirely different intervals.

The music itself tells you what the right objective function is. This shifts "musical context" from qualitative discussion to a measurable parameter.

Can mathematics tell us what sounds beautiful?

Maybe not. But it can tell us exactly what we are trading away, and that is worth knowing.

01

Tuning "optimality" is conditional. It requires an explicit objective function, target set, and weights.

02

Different norms and weightings produce genuinely different rankings of temperaments.

03

Repertoire can induce weights, connecting compositional intent to mathematical evaluation.

Limitations and natural extensions

Chord-level targets

Evaluate full chord impurity directly, not just individual intervals. A major triad is more than a third plus a fifth.

Psychoacoustic dissonance

Replace cents error with perceptual models accounting for timbre and critical-band interaction (Sethares, 2005).

Historical tuning vectors

Evaluate Werckmeister, Kirnberger, Vallotti under the same loss machinery for systematic comparison.

Corpus-scale experiments

Induce weights from harmonic datasets. Which temperament best fits Bach's harmonic language as a statistical model?